

**METHOD AND APPARATUS FOR TESTING LIP-SYNC OF DIGITAL
TELEVISION RECEIVER**

Technical Field

5 The present invention relates to lip synchronization test method and apparatus for a Digital TV (DTV) receivers based upon time-indexed audio and video signals.

10 In particular, the present invention relates to lip synchronization test method and apparatus capable of simply testing a regeneration time difference between audio and video signals of a DTV receiver based upon a video signal having Transient Effect Area Test Signals (TATS) inserted therein and an additional audio sub-channel signal having a frame index signal inserted therein.

15

Background Art

20 A digital broadcasting system executes entire processes based upon digital signals, and therefore can realize clear images and high sound quality by subdividing low-to-high tone sound signals. Accordingly, the application of the digital broadcasting system is getting spread rapidly and widely through various media such as terrestrial, satellite and cable TVs.

25 A test may be executed on the function and performance of a DTV receiver, which receives from the DTV broadcasting system, by using a signal from the broadcasting system. However, although an operational test may be executed based upon the signal from an actual broadcasting system, the DTV broadcasting system does not provide any signal by which a
30 special function of the receiver can be tested or the performance of the receiver can be maintained steady. As a consequence, a developer or manufacturer of DTV receivers typically establishes an imaginary broadcasting system in order to test the function and performance of a DTV receiver
35 via the imaginary broadcasting system.

Fig. 1 schematically illustrates a concept of a DTV test via an imaginary broadcasting system. A stream generator 101 generates an audio/video test stream to be used in a test on the function or performance of a DTV receiver, a channel encoder 102 executes channel-encoding to the test stream, a channel up-converter 103 modulates and up-converts the channel-encoded test stream into an RF level and then outputs the up-converted test stream in the form of an RF signal, and a DTV receiver or set-top box 104 receives the RF signal. Then, a measuring instrument 105 such as an oscilloscope 105 or an HDTV/PC monitor 106 is used to execute a test to the function or performance of the DTV receiver based upon the received test stream.

The length of the test stream used in the manufacture of the DTV receiver as above must be extended infinite as in an actual broadcasting. Also very wide types of test streams are needed for the function or performance test of the DTV receiver. This in turn generates a large number of channels together with enormous data. However, since a computer storage media for storing the DTV test stream has a limited capacity, stream length is minimized and a limited length of stream is repeatedly generated. In case that the limited length of test stream is repeatedly generated, it is required to correctly synchronize the generation time of a video or audio element stream (ES) with that of a transport stream (TS). That is, when the test stream is repeatedly generated, any inconsistency between the TS and the ES may induce audio or video discontinuity, which in turn may cause a DTV receiver to be misconceived defective even though it had passed a quality test in a manufacturing site. As a consequence, means for synchronize the ES with the TS have been needed in the art.

Further, when a DTV receiver decodes a received audio/video resource constituted of a digital video stream and then outputs (i.e., outputs sound and displays a video

image) through audio/video processing, any time difference may occur between audio and video signals thereby mismatching the video image and the sound. The mismatched sound and image may dissatisfy a user who watches a DTV set.

5 Such a process of synchronizing the audio signal with the video signal is referred to as "lip synchronization." Occasionally, means for executing lip synchronization (e.g., hardware and/or software) may be provided inside the DTV receiver. As a well-known technique for enabling the lip
10 synchronization, the DTV receiver uses the time information of the audio/video signals. The technique counts the time information of the audio/video signals to find whether an audio frame proceeds in advance of or behind a video frame, and if a time difference between the audio and video signals
15 exceeding a reference time interval is found at a specific time point, discards or repeats an audio frame at that time point, thereby executing the lip synchronization.

 However, even though a lip synchronization control function is mounted, a manufactured DTV receiver is subject
20 to a test with the foregoing test equipment in order to inspect whether audio and video streams of the DTV receiver are consistent with each other. The lip synchronization is a factor to be regarded essential in a DTV system, and tested by measuring a time difference between the audio and video
25 signals in the DTV receiver.

 For example, where the time difference between the audio and video signals is in the range from 25msec to 40msec, that is, the audio signal proceeds in advance of the video signal within 25msec or behind the video signal for at
30 least 40msec. Then, regenerated image and sound appear matching each other in this range, thereby providing more natural image and sound to the audience.

 Conventionally, in order to measure a time difference between audio and video signals in a DTV receiver, a lip
35 synchronization test is executed by using a video sequence

and an audio signal, which are specifically constructed. That is, the video signal consists of a black part, a white part and a short line part. Then, a lip synchronization test is executed by inserting a signal for video signal
5 synchronization into the short line part, and using a tone signal, which exponentially decreases at a synchronization position, as an audio test signal.

However, this method has a problem in that a video sequence and an audio sequence must have specific
10 constructions dedicated for the lip synchronization test. Therefore, the lip synchronization test for manufactured DTV receivers require additional devices and techniques, thereby causing inconvenient test operation and poor productivity.

15 **Disclosure of the Invention**

An object of the present invention is to provide lip synchronization test method and system for a Digital TV (DTV) receivers based upon time-indexed audio and video signals.

In particular, the present invention is aimed to
20 provide lip synchronization test method and system for DTV receivers capable of simply executing a lip synchronization test on DTV receivers by inserting time index signals into video and audio signals and then measuring the waveforms of the time-indexed audio and video signals with a measuring
25 device such as an oscilloscope.

Another object of the present invention is to provide lip synchronization test method and system for DTV receivers capable of simply executing a lip synchronization test by inserting time-indexed signals into video and audio signals
30 and then calculating an audio/video time difference based upon a time difference measured with a measuring device such as an oscilloscope and field/frame numbers obtained from the time indexes.

Still another object of the present invention is to
35 provide lip synchronization test method and system for DTV

receivers capable of simply executing a lip synchronization test by preparing audio and video resources, inserting frame index signals and Transient Effect Area Test Signals (TATS) into audio and video signals, respectively, compressing the frame index signals and the TATS into digital bit streams, decoding the digital audio and video streams in a DTV receiver, and then measuring the waveforms of the decoded digital audio and video streams with a device such as an oscilloscope.

According to an aspect of the present invention for realizing the objects, there is provided a lip synchronization test method for DTV receivers, the method comprising the steps of: generating a digital audio stream having a frame index inserted therein and a video stream having TATS inserted therein; comparing the waveform of the time-indexed audio signal with that of the time-indexed video signal; and measuring and calculating a time difference d , between the audio and video signals based upon the time indexes of the time-indexed audio and video signals.

According to another aspect of the present invention for realizing the objects, there is provided a lip synchronization test method of DTV receivers, the method comprising the steps of: generating digital audio and video streams each having a time index inserted therein; comparing the waveform of the time-indexed audio signal with that of the time-indexed video signal; and measuring and calculating a time difference d , between the audio and video signals based upon the time indexes of the time-indexed audio and video signals.

According to still another aspect of the present invention for realizing the objects, there is provided a lip synchronization test system for Digital TV (DTV) receivers comprising: means for detecting a time index signal contained in an audio signal to discriminate a corresponding audio frame number, and calculating a corresponding audio frame

time t_a from the audio frame number; means for detecting a time index signal contained in the video signal to discriminate a corresponding video field number and calculating a corresponding video field time t_v from the video field number; means for measuring a time difference between the audio and video signals; and means for calculating a lip synchronization time d , based upon the time difference between the audio and video signals, the corresponding audio frame time and the video field time.

Brief Description of the Drawings

Fig. 1 illustrates an imaginary broadcasting system for a DTV receiver test;

Fig. 2 is a flowchart of a DTV lip synchronization test method according to the present invention;

Fig. 3 schematically illustrates a DTV picture format adopted in the present invention;

Fig. 4 illustrates a TATS structure according to the present invention;

Fig. 5 illustrates a video sequence structure including TATS according to the present invention;

Fig. 6 illustrates an audio signal structure that is frame-indexed for a lip synchronization test according to the present invention;

Fig. 7 illustrates the relationship between time-indexed audio and video signals according to the present invention;

Fig. 8 illustrates an exemplary video frame including TATS according to the present invention;

Fig. 9 illustrates exemplary audio and video test waveforms of a DTV receiver for a lip synchronization test according to the present invention; and

Fig. 10 is a block diagram illustrating the structure of a lip synchronization test system for a DTV receiver according to the present invention.

Best Mode for Carrying Out the Invention

Hereinafter an embodiment of a lip synchronization test method according to the present invention will be described with reference to the accompanying drawings.

Fig. 2 is a flowchart of a DTV lip synchronization test method according to the present invention.

Audio and video resource generation step S201 regenerates audio and video signals. Through computer simulation, the audio signal creates an additional audio sub-channel, and the video signal creates a video pattern to be tested.

Time index inserting step S202 inserts frame indexes into the audio signal to form a time-indexed audio signal (refer to Fig. 6), and inserts field indexes into the video signal to form a time-indexed video signal (refer to Figs. 4 and 5). The structures of the time-indexed audio and video signals will be described later in more detail with reference to Figs. 4 to 8.

Next, Audio and video stream generation step S203 compresses the time-indexed audio and video signals to generate audio and video streams. Finally, the audio and video streams are multiplexed with an instrument such as an MPEG analyzer to form a transport stream.

The audio and video streams generated as above are inputted into a DTV receiver to be tested, and then decoded in an audio and video decoding step S204.

The audio/video decoding is executed according to a well-known technique of the art. That is, the DTV receiver receives and parses the transport stream, separately decodes the transport stream into audio and video signals, and then outputs sound and a video image by processing the audio and video signals.

Audio and video signal waveform measuring step S205 captures and measures the waveforms of the decoded audio and

video signals with a measuring device such as an oscilloscope. This measuring step S205 obtains a time difference t_{dav} between the audio and video signals.

Audio and video signal information calculating step S206 calculates an audio time t_a and a video time t_v based upon the time index information (waveforms) of the time-indexed audio and video signals. As will be described later in detail, the audio time t_a is obtained by calculating an audio frame number n from the time index information of the audio signal waveform and then calculating the audio frame number n with an audio frame time, and the video time t_v is obtained by calculating a video field number m corresponding to the audio frame number with a video field rate field/sec.

Next, audio and video time difference calculating step S207 calculates a final lip synchronization time d_l based upon the audio and video times, a time difference $t_a - t_v$ between the audio and video times and an initial value $t_{DTSoffset}$ of a Decoding Time Stamp (DTS). The DTS is used for decoding time information in a video Packetized Elementary Stream (PES) packet, and indicates a time value during which compressed video and audio streams are transmitted from a system decoder to an ES decoder.

The initial value of the DTS has an initial delay time in order to maintain the fullness of a buffer at a constant value. Since the video signal requires a larger buffer size, the video signal has a longer restoration time than the audio signal and thus has a delay time of about 0.2sec.

Of course, the initial value of the DTS can be varied according to stream transmission rate, frame rate, decoder buffer size and so on.

Based upon an n^{th} audio frame time t_a , an m^{th} video field time t_v , an audio and video time difference t_{dav} measured with a measuring device and a DTS initial value $t_{DTSoffset}$, the audio/video time difference t_{dav} may be obtained as $d_l = t_a - t_v - t_{dav} - t_{DTSoffset}$.

As described above, it is possible to execute a lip synchronization test by measuring the waveforms of the time-indexed audio and video signals with the measuring device such as an oscilloscope, and then obtaining the audio and video time difference based upon the measured time index information.

Next, in view of the audio time index, the video time indexes and the measured waveforms which have been described above, an exemplary lip synchronization test method will be described with reference to Figs. 3 to 9.

First, Fig. 3 schematically illustrates a 1920X1080 picture format. In this format, a display aperture 300 has a 1888X1062 size. A horizontal blank 301 is located in the left of the display aperture 300, a vertical blank 302 is located above the display aperture 300, Transparent Effect Areas (TAs) 303 (of 9 pixels) adjoin the top and bottom edges of the display aperture 300, and TAs 304 of 16 pixels adjoin the right and left edges of the display aperture 300.

According to the present invention, Transient Effect Area Test Signals (TATS) are inserted into the vertical TAs in every video frame to insert indexes about corresponding fields like Vertical Interval Test Signals (VITS) in an analogue TV system. That is, the TATS are inserted into first 8 lines of the video frame, in which the TATS have different values according to respective fields in order to express field/frame indexes.

Fig. 4 schematically illustrates the TATS. As shown in Fig. 4, video time indexes in the present invention are adapted to index corresponding video field numbers based upon 4 levels. That is, the indexing is executed in quadruple digits, in which a field index is expressed with an index waveform 401 having four levels of 0, 1, 2 and 3. Since two fields form one frame in case of an interlace mode, a field index of for example '0213₄' as seen in Fig. 4 indicates a 39th field and therefore the bottom field of a 19th frame.

Fig. 5 illustrates a video sequence 501 including the TATS. As can be seen from Fig. 5, one video frame consists of two video fields, and field indexes are graded in their order based upon 4 levels (quadruple digits) of 0, 1, 2 and 3 beginning from the first field.

Where one video frame is 33.3msec, one field will correspond to 16.67msec.

Fig. 6 illustrates time indexes to be inserted into audio frames according to the present invention. As shown in Fig. 6, an index waveform is inserted into every audio frame for the purpose of frame indexing. That is, the time indexes are expressed as additional audio sub-channels each containing a frame index related with frame number.

In view of the exemplary audio frame time indexes shown in Fig. 6, one frame consists of 32msec, and an audio frame index signal graded (inserted) into each frame includes time information 601 and the number of signals 602.

For example, a first audio frame has a sinusoidal wave of one cycle inserted therein after 1msec, and a second audio frame has a sinusoidal wave of two cycles inserted therein after 2 msec.

Accordingly, the relationship between the video and audio time indexes described hereinbefore will be arranged as in Table 1 below.

Table 1

Video field number	TATS (Time index)	Audio frame number	Time index	Number of index signal
1	0000	1	1msec	1 cycles
2	0001	2	1msec	2 cycles
3	0002
4	0003	
5	0010	10	1msec	10 cycles
....		11	2msec	1 cycles
21	0110	12	2msec	2 cycles
....	

120	1313
....		90	9msec	10 cycles
164	2210	91	10msec	1 cycles
....	
256	3333	125	12msec	5 cycles

The video field and audio frame indexes as above are looped back for a specific time period. For example, as shown in Fig. 7, a video frame number and an audio frame number 702 are looped back every 4 seconds in order to avoid any confusion between adjacent synchronous signals.

Fig. 8 illustrates an exemplary video frame having TATS inserted therein according to the present invention. It is to be understood that time indexes, that is, field indexes as described above are inserted into a TA 801. Since it is assumed that one frame is constituted of two fields such as top and bottom fields as described above, the field index is expressed as a top field index 802 and a bottom field index 803. In Fig. 8, since a TATS level in a TA has a value of '3230₄' when it is a top field, it is understood that this value means a 236th field and therefore expresses a 118th frame index.

When the waveforms of time-indexed video and audio signals as reported in Table 1 above are observed with a measuring instrument such as an oscilloscope, an audio time t_a and a video time t_v can be calculated based upon the frame number of audio signals and the field (frame) number of video signals, and respective frame rates or frame times thereof can be calculated from the time indexes. Then, audio and video lip synchronization times can be calculated by using the audio and video times t_a and t_v .

Fig. 9 illustrates exemplary audio and video test waveforms of a DTV receiver for a lip synchronization test according to the present invention. The waveforms in Fig. 9 are captured measured with an oscilloscope, in which the

reference numeral 901 represents an audio signal, and the reference numeral 902 represents a video signal.

Since a sinusoidal wave of one cycle is detected at 10msec in the audio signal 901, it can be understood that this time index signal indicates a 91st audio frame as reported in Table 1 above. Therefore, the time t_a of the 91st audio frame is calculated $t_a = 32(\text{msec/frame}) \times 91(\text{frame}) = 2.912\text{sec}$ according to an audio frame time 23msec.

In addition, a time index (TATS) waveform included in the video signal 902 has a value '2210₄', which indicates a 164th field, that is, an 82nd frame. Therefore, the time t_v of the 164th video field is calculated $t_v = 164/60 = 2.7333\text{sec}$ according to the field rate 60 field/sec.

Since it can be understood that the time difference t_{dav} measured between the 91st audio frame and the 164th video field is 2msec, the lip synchronization time, that is, the time difference between the audio and video signals is calculated

$$d_t = t_a - t_v - t_{dav} - t_{DTSoffset} = 2.912 - 2.733 - 0.002 - 0.2 = -0.023\text{sec}$$

when a DTS initial value $t_{DTSoffset}$ is set 0.2sec.

This test result teaches that the audio signal is behind the video signal for 23msec and therefore within a lip synchronization standard (+25msec to -40msec). This means that the corresponding DTV receiver passes the lip synchronization test.

Fig. 10 illustrates the structure of a lip synchronization test system for DTV receivers according to the present invention, which is designed to automatically execute the lip synchronization test method for DTV receivers as described hereinbefore.

The lip synchronization test system for DTV receivers according to the present invention includes an audio time calculating means 11 for detecting a time index signal contained in an audio signal to discriminate a corresponding

audio frame number, and calculating a corresponding audio frame time t_a from the audio frame number; a video time calculating means 12 for detecting a time index signal contained in the video signal to discriminate a corresponding video field number and calculating a corresponding video field time t_v from the video field number; a measuring means 13 for measuring a time difference between the audio and video signals; and a lip synchronization discriminating means 14 for calculating a lip synchronization time d , based upon the time difference between the audio and video signals, the corresponding audio frame time and the video field time.

The audio time calculating means 11 includes an audio time index detecting means 111 for detecting a time index signal contained in the audio signal, an audio frame number discriminating means 112 for decoding the detected audio time index signal to discriminate the corresponding audio frame number and a calculating means 113 for calculating the discriminated audio frame number with an audio one frame time to calculate the corresponding audio frame time t_a .

The video time calculating means 12 includes a video time index detecting means 121 for detecting a time index signal contained in the video signal, a video field number discriminating means 122 for decoding the detected video time index signal to discriminate the corresponding video field number and a calculating means for calculating the discriminated video field number with a video field rate to obtain the corresponding video frame time t_v .

The operation of the lip synchronization test system for DTV receivers of the present invention of the above structure will be described as follows.

The audio time index detecting means 111 detects an inserted audio time index signal as shown in Fig. 6 from an input audio signal. The audio time index detecting means 111 detects the audio time index signal that is inserted into each audio sub-channel in the form of a sinusoidal wave,

which has n number of cycles and a predetermined time in each audio frame.

The audio frame number discriminating means 112 counts a time period from an initial position of the audio frame until a sinusoidal wave appears and counts the number of detected sinusoidal waves to discriminate a corresponding audio frame. For example, since a sinusoidal wave of one cycle is counted after 10msec in the waveform shown in Fig. 9, a 91st audio frame can be decoded from this value.

By using the decoded audio frame number n and an audio one frame time of for example 32msec, the calculating means calculates a corresponding audio frame time t_a expressed as $t_a = 32\text{msec} \times n[\text{sec}]$ as described above, and inputs the calculated audio frame time into the lip synchronization discriminating means 14.

The video time index detecting means 121 detects a video time index signal, which is inserted as shown in Figs. 4 and 5, from an input audio signal. Since the video time index signal in the form of 4 levels is inserted into each video field, a signal having 4 levels is detected as the video time index signal. Herein, by detecting the video time index corresponding to the audio frame, the lip synchronization time difference can be measured correctly.

The video field number discriminating means 122 decodes the level of the detected video time index signal to discriminate a corresponding video field number. For example, a 164th video field (82nd video frame) can be decoded from the value '2210₄' in the waveform shown in Fig. 9.

The calculating means 123 calculates a corresponding video field time $t_v = m/60 [\text{sec}]$ by using the decoded video field number m and a video field rate of for example 60 field/sec, and inputs the calculated video field time into the lip synchronization discriminating means 14.

The measuring means 13 measures a time difference between the audio frame calculated by the audio time

calculating means 11 and the video frame calculated by the video time calculating means 12. As an example of a measuring method, the measuring means 13 can count a time (e.g., 2msec in Fig. 9) from an initial point of an audio frame until detecting a corresponding field index as shown in Fig. 9 in order to measure the audio and video time difference t_{dav} .

Based upon the n^{th} audio frame time, the m^{th} video field time corresponding to the n^{th} audio frame time, the detected audio and video time difference t_{dav} and the DTS initial value $t_{DTSoffset}$, the lip synchronization discriminating means 14 calculates a lip synchronization time $d_t = t_a - t_v - t_{dav} - t_{DTSoffset}$, and outputs a calculated result in the form of numeric values. Alternatively, the lip synchronization discriminating means 14 can judge whether a corresponding DTV receiver is defective or not by comparing the result with a predetermined reference range.

Industrial Applicability

The present invention can very simply execute a lip synchronization test on a DTV receiver by capturing the waveforms of time-indexed audio and video signals to obtain the time information of the audio and video signals and execute the lip synchronization test based upon the measured time difference between the audio and video signals.

Further, the present invention advantageously executes a lip synchronization test by using video TATS and additional audio sub-channels without any additional lip synchronization test program.

In particular, the lip synchronization test method for DTV receivers according to the present invention has an advantage in that a lip synchronization test can be executed on the waveforms of time-indexed audio and video signals regardless of time points when the waveforms of time-indexed audio and video signals are captured.